SPECIAL
THORPEX
COLLECTION

Characteristics of Target Areas Selected by the Ensemble Transform Kalman Filter for Medium-Range Forecasts of High-Impact Winter Weather

SHARANYA J. MAJUMDAR AND KATHRYN J. SELLWOOD

Rosenstiel School of Marine and Atmospheric Science, Miami, Florida

DANIEL HODYSS

Naval Research Laboratory, Monterey, California

ZOLTAN TOTH AND YUCHENG SONG

NOAA/NCEP Environmental Modeling Center, Camp Springs, Maryland

ABSTRACT

The characteristics of "target" locations of tropospheric wind and temperature identified by a modified version of the ensemble transform Kalman filter (ETKF), in order to reduce 0–7-day forecast errors over North America, are explored from the perspective of a field program planner. Twenty cases of potential high-impact weather over the continent were investigated, using a 145-member ensemble comprising perturbations from NCEP, ECMWF, and the Canadian Meteorological Centre (CMC).

Multiple targets were found to exist in the midlatitude storm track. In half of the cases, distinctive targets could be traced upstream near Japan at lead times of 4–7 days. In these cases, the flow was predominantly zonal and a coherent Rossby wave packet was present over the northern Pacific Ocean. The targets at the longest lead times were often located within propagating areas of baroclinic energy conversion far upstream. As the lead time was reduced, these targets were found to diminish in importance, with downstream targets corresponding to a separate synoptic system gaining in prominence. This shift in optimal targets is sometimes consistent with the radiation of ageostrophic geopotential fluxes and transfer of eddy kinetic energy downstream, associated with downstream baroclinic development. Concurrently, multiple targets arise due to spurious long-distance correlations in the ETKF. The targets were least coherent in blocked flows, in which the ETKF is known to be least reliable. The effectiveness of targeting in the medium range requires evaluation, using data such as those collected during the winter phase of The Observing System Research and Predictability Experiment (THORPEX) Pacific Asian Regional Field Campaign (T-PARC) in 2009.

1. Introduction

The global atmospheric observational network has traditionally comprised land-based rawinsonde balloons and satellite-borne sensors. Yet, as stated by Lorenz and Emanuel (1998), " . . . despite this wealth of data—more, in fact, than we know how to use to full advantage—large gaps remain in our picture of the global weather pattern, particularly over the less frequently visited areas of the oceans." In an effort to fill these gaps, several field campaigns have taken place over the past decade, in which the routine observational network has been augmented by a limited number of *adaptive* observations. These observations have been "targeted" at improving short-range (1–3 day) numerical forecasts of weather events such as frontal cyclones and hurricanes (see Langland 2005 for a review). The instrumentation utilized to date has primarily been Global Positioning System (GPS) dropwindsondes, released from manned aircraft deployed over the oceans. The assimilation of these targeted observations has on average improved short-range

Corresponding author address: Dr. Sharanya J. Majumdar, Division of Meteorology and Physical Oceanography, RSMAS, Miami, FL 33149.
E-mail: smajumdar@rsmas.miami.edu

Report Documentation Page

1. REPORT DATE **24 NOV 2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Characteristics of Target Areas Selected by the Ensemble Transform Kalman Filter for Medium-Range Forecasts of High-Impact Winter Weather**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Monterey,CA,93943**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The characteristics of ??target?? locations of tropospheric wind and temperature identified by a modified version of the ensemble transform Kalman filter (ETKF), in order to reduce 0?7-day forecast errors over North America, are explored from the perspective of a field program planner. Twenty cases of potential highimpact weather over the continent were investigated, using a 145-member ensemble comprising perturbations from NCEP, ECMWF, and the Canadian Meteorological Centre (CMC). Multiple targets were found to exist in the midlatitude storm track. In half of the cases, distinctive targets could be traced upstream near Japan at lead times of 4?7 days. In these cases, the flow was predominantly zonal and a coherent Rossby wave packet was present over the northern Pacific Ocean. The targets at the longest lead times were often located within propagating areas of baroclinic energy conversion far upstream. As the lead time was reduced, these targets were found to diminish in importance, with downstream targets corresponding to a separate synoptic system gaining in prominence. This shift in optimal targets is sometimes consistent with the radiation of ageostrophic geopotential fluxes and transfer of eddy kinetic energy downstream, associated with downstream baroclinic development. Concurrently, multiple targets arise due to spurious long-distance correlations in the ETKF. The targets were least coherent in blocked flows, in which the ETKF is known to be least reliable. The effectiveness of targeting in the medium range requires evaluation, using data such as those collected during the winter phase of The Observing System Research and Predictability Experiment (THORPEX) Pacific Asian Regional Field Campaign (T-PARC) in 2009.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**22**	

forecast skill (Szunyogh et al. 2002; Aberson 2003). Several of these field campaigns have relied on guidance from mathematical *adaptive observing strategies*, which identify locations for adaptive sampling that are "optimal" in that observations in these locations are expected to most significantly reduce errors in some selected aspect of the forecast.

One primary goal of the World Meteorological Organization (WMO)-sanctioned The Observing System Research and Predictability Experiment (THORPEX)[1] program is to improve medium-range (3–7 days) forecasts of high-impact weather, via intelligent assimilation of targeted observations from novel platforms, based on a reliable adaptive observing strategy (Rabier et al. 2008). New types of observations may comprise, for example, zero-pressure driftsonde balloons releasing dropwindsondes along the storm track, satellite-based atmospheric motion vectors in rapid-scan mode (Velden et al. 2005), wind and relative humidity profiles from airborne lidar (Weissmann et al. 2005), or dropwindsondes and remotely sensed observations mounted aboard high-altitude, long-endurance unmanned aircraft systems (MacDonald 2005). In contrast to campaigns such as Winter Storm Reconnaissance (WSR; Szunyogh et al. 2002) in which limited targeted data are collected at one synoptic time, one may consider temporally *continuous* data from satellites or unmanned platforms over the ocean. For instance, in order to improve 3–7-day forecasts of a winter cyclone that impacts North America, the question arises as to whether a platform may be able to sample continuously and cost effectively in propagating locations over the northern Pacific Ocean. This paper represents an initial attempt to identify and understand the properties of locations deemed optimal for adaptive sampling in this context. More specifically, we focus on the following question: If a high-impact weather event is anticipated over North America over a week in advance, are we able to identify dynamically coherent upstream regions in which supplementary observations would benefit the medium-range forecast?

The adaptive sampling strategy of choice in this paper is the ensemble transform Kalman filter (ETKF; Bishop et al. 2001), which predicts the reduction in *forecast error variance* within a given *verification region* at time t_v due to the assimilation of any hypothetical group of observations at an earlier analysis time t_a. Forecast error variance, or forecast uncertainty, may be predicted by a *multimodel* ensemble of forecasts, aimed at capturing a wide range of likely scenarios. The ETKF is philosophically different from adjoint-based methods in this regard, since the latter methods (e.g., Palmer et al. 1998) are aimed at reducing the forecast error in a *given model*. On the other hand, theoretical connections between ensemble- and adjoint-based techniques have been identified (Leutbecher 2003; Majumdar et al. 2006; Ancell and Hakim 2007). The ETKF has been used in the National Oceanic and Atmospheric Administration (NOAA) operations to plan WSR missions, and it has proven to be an effective technique within the short range (Majumdar et al. 2001, 2002). The optimal target area for forecasts is often a single region in which strong mid- and upper-tropospheric winds exist, or a region of baroclinic instability (Petersen et al. 2007). However, in the medium range, the picture is expected to be less clear, since one may expect multiple locations for adaptive sampling due to the dispersive nature of Rossby waves in the midlatitudes, and possible influences on midlatitude weather by tropical or polar phenomena. Moreover, the ETKF is only perfect when linear dynamics hold and the model and error covariance specification are perfect. Cognizant of these caveats, Sellwood et al. (2008) performed an initial investigation into the capability of the ETKF to predict the influence of WSR dropwindsonde data released over the northeastern Pacific Ocean on medium-range forecasts downstream. Using a 51-member European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble (Buizza et al. 2003), they concluded that the ETKF is capable of discriminating between observation locations that are effective and ineffective for 3–6-day National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) forecasts of 200-hPa winds within a verification region based on Rossby wave dispersion, if the flow was predominantly zonal. The ETKF performed poorly in blocked flows. The results of Sellwood et al. (2008) were encouraging since they were based on an ensemble of limited size, the flow beyond 2 days was influenced by the North American continent and therefore often nonzonal, and the influence of the assimilation of the targeted data on the NCEP GFS forecast was contaminated by initially remote, small-scale noise that rapidly grew and muddled the dynamical influence from the observation locations (Hodyss and Majumdar 2007). In this paper, we use a multimodel ensemble comprising 145 forecasts from three operational centers. We therefore expect the covariance structure to be more accurate than that derived by Sellwood et al. (2008), given that the uncertainty derived from a multimodel ensemble is expected to be closer to the true range of possibilities than that represented by a single model ensemble. It is worth noting, however, that correlated model error may become important at longer lead times. We emphasize that while the ETKF combines data assimilation with predictions of evolving error variance (Bishop et al. 2001), this paper

---

[1] "THORPEX: A World Weather Research Programme" is a component of the WMO World Weather Research Programme.

focuses on the evolution of *observation sensitivity*. In other words, this paper aims to examine qualitatively the characteristics of the targets, as opposed to the practical application in which actual targeted data are assimilated. Our goal is to provide hypotheses for field campaigns in which the actual effects of assimilating observations in these target areas may be evaluated quantitatively.

A starting point for developing such hypotheses for the temporal evolution of the ETKF targets is to consider the conclusion of Szunyogh et al. (2002), who claimed that the propagation of the effect of the targeted observations is associated with *downstream baroclinic development*. If this claim holds, one might expect that target regions would evolve downstream in a similar manner as the forecast lead time $(t_v - t_a)$ decreases from 7 to 0 days (where t_v is fixed). Several decades after observations of downstream development of baroclinic waves in the middle and upper troposphere were documented (beginning with Namias and Clapp 1944), numerical investigations determined that downstream baroclinic development was achieved by examining the response of a baroclinically unstable atmosphere to a local initial perturbation (Simmons and Hoskins 1979). A series of papers in the 1990s investigated this phenomenon in greater depth, via theoretical and observational studies of the eddy kinetic energy budget (Orlanski and Katzfey 1991; Orlanski and Sheldon 1993, 1995; Orlanski and Chang 1993; Chang 1993; Chang and Orlanski 1993, 1994). These papers argued that a wave in the storm track may first amplify rapidly via baroclinic energy conversion until it reaches maturity, after which its center begins to weaken as energy is radiated downstream via ageostrophic geopotential flux divergence, amplifying the downstream perturbations. This process may continue, with this second energetic area eventually weakening as another new center develops farther downstream. Chang (1993) indicated that baroclinic energy conversion is only of secondary importance in the eastern Pacific, as the wave approaches North America. Danielson et al. (2004) claimed that the connection between eddy energy in the zone of warm ascent in western North Pacific cyclones and cold cyclones in the eastern North Pacific is partially explained by downstream baroclinic development. Processes such as stationary waves initiated by orographic forcing or zonal variations in diabatic heating (e.g., Hoskins and Valdes 1990) may act to decrease eddy activity and complicate the picture of downstream baroclinic development over the continent. In summary, baroclinic energy conversion is hypothesized to dominate upstream, in locations over the western Pacific Ocean off Japan, while the storm track is extended in the central Pacific via radiating geopotential fluxes in a relatively stable area. We will compare the ETKF targets with these processes that form the primary forcing terms in the eddy kinetic energy equation.

Our investigation is based on 20 cases of extratropical high-impact weather during January–March 2007, in the context of a field program planner who is required to make a decision on deployments *at least 7 days in advance* of the anticipated weather event. As the lead time $(t_v - t_a)$ is increased from 0 to 7 days, the optimal target areas are expected to track upstream from the verification region, perhaps with a connection to the process of downstream baroclinic development. In section 2, a modified version of the ETKF, based on constraining the distribution and magnitude of ensemble-based analysis error variance to reflect that of an operational data assimilation scheme, is presented. This version of the ETKF was transitioned for use at NCEP during the winter phase of the THORPEX Pacific Asian Regional Campaign (T-PARC) in 2009. Some characteristics of the target regions are illustrated for two cases in section 3. A summary of the results over all cases and some general characteristics are provided in section 4, followed by the conclusions in section 5.

2. Experimental design

a. Case selection

Twenty cases of high-impact weather over North America were selected (Table 1). The verification times were coincident with weather systems that were judged by forecasters at NOAA's Hydrometeorological Prediction Center (HPC) to be of medium or high priority during the 2007 WSR Program.[2] The weather systems were synoptically driven, comprising landfalling cyclones on the West Coast, winter storms in the east, and rain and snow events in the central states. Cases at least 2 days apart were chosen. For consistency of comparison between the 20 cases, a fixed "verification region" encompassing a large area of North America (25°–65°N, 125°–75°W) was selected. The ETKF guidance is based on ensemble forecasts initialized 10 days prior to the verification time. While the ensemble may not have been able to capture the accurate timing, location, and strength of the weather event that was identified by HPC forecasters 3–5 days in advance, several ensemble members contained the relevant synoptic-scale features within the verification region, such as a deep 500-hPa trough. An ensemble initiated 10 days prior to the verification time was chosen in order to account for the lead time required

---

[2] Case 1, in which major ice storms affected much of the central and southern United States, was also selected in this study because of its significance, even though WSR 2007 had not yet begun.

TABLE 1. List of 20 cases used in this study, their forecast verification time, and the predicted weather event and priority assigned by NOAA/HPC forecasters.

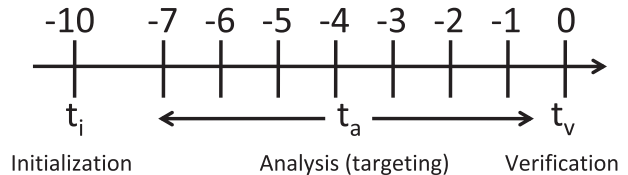| Case	Verification time	Description of weather event(s)	Priority
1	15 Jan	Ice storm over southern United States	N/A
2	19 Jan	CA low; MS valley rain; East Coast cyclogenesis	Medium
3	22 Jan	Southern stream low	Medium
4	25 Jan	Potential snow in Northeast	Medium
5	27 Jan	Central West Coast precipitation	Medium
6	31 Jan	Lower MS valley rain	High
7	2 Feb	Gulf Coast rain; Southeastern U.S. low	Medium
8	5 Feb	Mid-Atlantic coast snow potential	High
9	9 Feb	Heavy rain/snow in northern CA	Medium
10	14 Feb	East Coast winter storms	High
11	16 Feb	Precipitation in CA	Medium
12	19 Feb	Storms in NV	High
13	21 Feb	MS valley rain	High
14	23 Feb	Precipitation in CA	Medium
15	25 Feb	Storm in plains; central and eastern storms	High
16	27 Feb	Pacific Northwest precipitation	Medium
17	2 Mar	Rain in Appalachians	High
18	3 Mar	Pacific NW precipitation	Medium
19	11 Mar	Pacific NW precipitation; MS valley severe weather	High
20	15 Mar	Midwest precipitation	Medium

FIG. 1. Timeline for targeting used in this paper. Ensembles initialized 10 days prior to the verification time are used to provide summary maps of signal variance, based on targeting at any of the sequence of times 0–7 days prior to the verification time.

also reduces any undesirable contributions from the variance of the means in the covariance calculation. In contrast, our motivation for computing an ensemble perturbation about the mean of the ensemble from which that member is drawn (instead of the multiensemble mean) is to produce more realistic samples from the true distribution of dynamically evolving initial condition errors. Using this method, we expect to reduce the likelihood of these initial condition perturbations being dominated by systematic differences between the NCEP, ECMWF, and CMC models, which may not resemble the dynamically evolving structures in ensembles described by Buizza and Palmer (1995), Toth and Kalnay (1997), and Magnusson et al. (2008).

The modified version of the ETKF code [i.e., "Variational" ensemble transform Kalman filter (Var-ETKF)], used during the summer and winter phases of T-PARC during 2008–09, is now presented. The ETKF uses the ensemble perturbations to predict the reduction in forecast error variance within a given verification region at verification time t_v, based on the assimilation of adaptive observations at analysis time t_a ($t_i < t_a < t_v$). For each case in this study, the ensemble initialization time t_i is held fixed, and $t_v − t_a$ is varied between 0 and 7 days (Fig. 1). In all ensemble filters, an error covariance matrix is represented as

$$\mathbf{P}^i(t|H^i) = \mathbf{Z}^i(t|H^i)\mathbf{Z}^{i\mathrm{T}}(t|H^i). \qquad (1)$$

Serial assimilation theory (Bishop et al. 2001) is then used to break the observational network at time t_a into "routine" (operator $\mathbf{H}^r$, error covariance $\mathbf{R}^r$) and "adaptive" (operator $\mathbf{H}^q$, error covariance $\mathbf{R}^q$) components. Analysis and forecast error covariance matrices $\mathbf{P}(t|\mathbf{H})$ are then computed via a linear transformation $\mathbf{T}$ of the ensemble that is based on these subsets of the observational network. The previous version of the ETKF used in Majumdar et al. (2006) and Sellwood et al. (2008), and also operationally during WSR, had assumed a crude simplification of the routine observational network, only accounting for rawinsondes and sparsely distributed brightness temperatures. The ETKF was then used to

for the planning of a long-endurance unmanned aircraft mission, the launching of a driftsonde balloon, or the activation of rapid-scan mode aboard satellites.

b. "Variational" ensemble transform Kalman filter

A combined ensemble of NCEP GFS,[3] ECMWF, and Canadian Meteorological Centre (CMC) forecasts, using 60, 51, and 34 available daily members from the respective ensembles (Toth and Kalnay 1997; Buizza et al. 2003; Pellerin et al. 2003), is used to prepare a matrix $\mathbf{Z}^i(t|H^i)$ that contains 144 linearly independent forecast perturbations, initialized at time t_i with operator H^i. The perturbation for each ensemble member is computed about the mean of the ensemble from which that member is drawn. We note that the optimal method to compute perturbations from a multimodel ensemble has not been formally established. An argument may be made for computing the perturbations about a multiensemble mean, which represents the minimum error variance estimate and

---

[3] For case 8, the full NCEP GFS ensemble was not available. A combined ECMWF and CMC ensemble was used for this case.

transform the raw ensemble $\mathbf{Z}^i(t_a|H^i)$ via the matrix $\mathbf{T}^r$ to produce the routine analysis error covariance matrix valid at time t_a:

$$\mathbf{P}^r(t_a|\mathbf{H}^r) = \mathbf{P}^i(t_a|H^i) - \mathbf{P}^i(t_a|H^i)$$
$$\times \mathbf{H}^{r\mathrm{T}}[\mathbf{H}^r\mathbf{P}^i(t_a|H^i)\mathbf{H}^{r\mathrm{T}} + \mathbf{R}^r]^{-1}\mathbf{H}^r\mathbf{P}^i(t_a|H^i), \tag{2a}$$

$$= \mathbf{Z}^i(t_a|H^i)\mathbf{T}^r\mathbf{T}^{r\mathrm{T}}\mathbf{Z}^{i\mathrm{T}}(t_a|H^i). \tag{2b}$$

To produce the ETKF guidance maps, the same equations are first solved to compute the analysis error variance due to the qth deployment of targeted observations:

$$\mathbf{P}^{r+q}(t_a|\mathbf{H}^{r+q}) = \mathbf{P}^r(t_a|\mathbf{H}^r) - \mathbf{P}^r(t_a|\mathbf{H}^r)\mathbf{H}^{q\mathrm{T}}$$
$$\times [\mathbf{H}^q\mathbf{P}^r(t_a|\mathbf{H}^r)\mathbf{H}^{q\mathrm{T}} + \mathbf{R}^q]^{-1}\mathbf{H}^q\mathbf{P}^r(t_a|\mathbf{H}^r). \tag{3}$$

This equation is then extended to time t_v to yield the associated forecast error covariance $\mathbf{P}^{r+q}(t_v|\mathbf{H}^{r+q}) = \mathbf{P}^r(t_v|\mathbf{H}^r) - \mathbf{S}^q(t_v|\mathbf{H}^q)$, where

$$\mathbf{S}^q(t_v|\mathbf{H}^q) = \mathbf{Z}^i(t_v|\mathbf{H}^r)\mathbf{Z}^{i\mathrm{T}}(t_a|\mathbf{H}^r)\mathbf{H}^{q\mathrm{T}}[\mathbf{H}^q\mathbf{P}^r(t_a|\mathbf{H}^r)\mathbf{H}^{q\mathrm{T}}$$
$$+ \mathbf{R}^q]^{-1}\mathbf{H}^q\mathbf{Z}^i(t_a|\mathbf{H}^r)\mathbf{Z}^{i\mathrm{T}}(t_v|\mathbf{H}^r). \tag{4}$$

Further details on the ensemble transformations that enable fast computations of this product are given in Majumdar et al. (2002). The diagonal of $\mathbf{S}^q(t_v|\mathbf{H}^q)$ localized within the verification region is then plotted as a function of the qth targeted observation on the ETKF guidance used throughout this paper. The qth targeted observation is a "pseudo sounding" of (u, v, T) at the 200-, 500-, and 850-mb levels, respectively, sampled at a model grid point at 2° resolution. The summary map therefore represents a mosaic of reduction in forecast error variance within the fixed verification region, as a function of the observation location. The optimal target location is the value of q for which $\mathbf{S}^q(t_v|\mathbf{H}^q)$ is largest.

We found that the geographical distribution of routine analysis error variance defined in (2a) did not resemble an operational distribution of analysis error variance. Instead, it possessed very large values of variance along features of high gradient over the oceans, and tiny values over land (see Fig. 8 of Reynolds et al. 2007). We therefore decided that a more realistic estimate of the ensemble-based routine analysis error variance was necessary. Reverting to Majumdar et al. (2002), an ensemble transform (ET) was used together with the monthly mean root-mean-square analysis error $\mathbf{P}^i_{\mathrm{OPER}}(t_a|H^i)$ provided by the Naval Research Laboratory (NRL) Atmospheric Variational Data Assimilation System (NAVDAS) to solve for $\mathbf{T}^r$, thereby replacing (2a) with

$$\mathbf{T}^{r\mathrm{T}}\mathbf{Z}^i[\mathbf{P}^i_{\mathrm{OPER}}(t_a|H^i)]^{-1}\mathbf{Z}^i\mathbf{T}^r = \mathbf{I}. \tag{5}$$

This direct transformation rotates and rescales the ensemble perturbations in order to provide a global distribution of analysis error variance that is more consistent with that produced by an operational data assimilation scheme. The full state-dependent ensemble information is still maintained in the transformed ensemble. For these calculations, no information on intervening observational networks between times t_i and t_a is required, while it is necessary in (2). It was found that the ET transformation in (5) produced a more even distribution of routine analysis error variance in the transformed ensemble than that produced by (2), similar to the aforementioned result of Reynolds et al. (2007). For similar reasons, the ET has been extended for ensemble generation by McLay et al. (2008) in the U.S. Navy ensemble prediction system, and by Wei et al. (2008) for the NCEP ensemble. We refer to this new version of the ETKF as the Var-ETKF [analogous to variational singular vectors (Var-SVs) introduced by Gelaro et al. (2002) and used by Reynolds et al. (2007)]. The Var-ETKF produces stronger sensitivity over land, compared with the older version that erroneously overstates the oceanic regions as target locations due to falsely large ensemble-based analysis errors. Throughout this paper, the value of $\mathbf{T}^r$ computed using (5) is used in the computation of reduction of forecast error variance in (4).

Theoretically, the reduction in forecast error variance is equal to the variance of "signals," where a signal is defined as the difference between two numerical forecasts that are identical in every respect except that one forecast includes the targeted data in the assimilation, while the second withholds the targeted data (Bishop et al. 2001). Therefore, the ETKF attempts to predict the variance of the propagation of the effect any set of targeted observations, and it is this quantity that was evaluated first by Majumdar et al. (2001, 2002) and more recently by Sellwood et al. (2008), who showed that the ETKF is able to predict signal variance adequately for 0–6-day forecasts in predominantly zonal flow regimes, but not in blocked flows. Unlike Sellwood et al. (2008), who used a 200-hPa wind norm, we use a vertically averaged difference total energy (DTE) norm throughout this paper:

$$E(x, y, t) = \frac{1}{p_t - p_0}\int_{p_0}^{p_t}\left[(u_s^2 + v_s^2) + \frac{c_p}{T_r}T_s^2\right]\frac{dp}{2}. \tag{6}$$

The quantities u_s, v_s, and T_s represent the signal in the zonal and meridional wind, and temperature fields, respectively. The vertical average is computed over three

pressure levels: 850, 500, and 200 hPa, and T_r is a reference temperature (287 K). We elect to use the DTE norm for consistency with previous papers on predictability and targeted observations (Palmer et al. 1998; Buizza and Montani 1999; Majumdar et al. 2002; Zhang et al. 2003; Petersen et al. 2007). Other perturbation variables such as moisture were found to produce a secondary contribution to the signal variance.

c. Eddy kinetic energy diagnostics

The vertically averaged form of the equation governing the evolution of eddy kinetic energy (EKE) follows from Chang (1993):

$$\frac{\partial K_e}{\partial t} + \mathbf{V}_m \cdot \nabla K_e + \upsilon \cdot \nabla_3 K_e = -(\nabla \cdot \upsilon \varphi) - \omega \alpha$$
$$+ \text{ minor terms.} \quad (7)$$

The EKE K_e is the kinetic energy of the eddy component υ of the velocity field, computed about a monthly mean $\mathbf{V}_m$. The vertical average is computed between 1000 and 100 hPa, at 100-hPa intervals in this paper. The terms on the left-hand side represent the local EKE tendency, the horizontal advection of EKE by the mean flow, and the advection by the eddy component of the flow. The first two terms on the right-hand side are the *ageostrophic geopotential flux convergence* and the *baroclinic energy conversion*, respectively. These two quantities have been found in this study to dominate over the remaining forcing terms such as barotropic energy conversion or eddy dissipation of energy. In this paper, we concentrate on diagnosing the relationship (if any) between ETKF targets and EKE maxima, and its creation and reduction via baroclinic energy conversion and radiation of ageostrophic geopotential fluxes. We use NCEP GFS forecasts initialized 7 days prior to the verification time to compute the quantities in (7), since the model output at 1° resolution was only available out to 7 days.

3. Examples

a. Case 15: Major snowstorm

First, the 15th of the 20 cases is chosen to illustrate some typical characteristics of the ETKF guidance over the 0–7-day forecast lead times. An unusually cold Pacific storm system made landfall on the coast of central California on 22 February 2007, producing up to 3 ft of snow at higher elevations. High winds, thunderstorms, and hail accompanied the storm system. The system propagated into the central high plains and evolved into a major snowstorm, with blizzard or winter storm warnings issued in 15 states. The high eddy kinetic energy and

baroclinic energy conversion associated with the storm over the central United States at the verification time (0000 UTC 25 February 2007) is evident in the 7-day NCEP GFS forecast (Fig. 2h).

The evolution of the ETKF targets as the sampling time t_a approaches the verification time t_v is illustrated in Fig. 2. At a lead time of −7 days (Fig. 2a), two main targets exist: the first (labeled A) being offshore of southern Japan in an area of baroclinic energy conversion, and the second (labeled B) along an elongated area of high EKE downstream. At −6 days, the targets have shifted ∼15° toward the verification region. Target B has shifted emphasis toward the northeastern extent, collocated with the EKE maximum. However, B diminishes in importance shortly thereafter, while A becomes dominant at −5 days and is situated downstream of an elongated trough. The ETKF sensitivity associated with the northern side of the trough near Kamchatka diminishes over −4 and −3 days. In contrast, the southern branch of the trough propagates eastward, and the associated ETKF sensitivity appropriately lies in the baroclinic zone downstream, east of the date line at −3 days. At −2 and −1 days, the sensitivity appears elongated, particularly across the ridge where there exists a broad distribution of EKE. The cutting off of the upper-level trough and renewed amplification of energy over the central United States between −1 day and the verification time (0 days) led to the widespread winter storms inland. As expected, the optimal target location with no lead time was this storm system itself. In summary, there is spatiotemporal continuity associated with the targets labeled A over a 7-day period, extending from Japan at −7 days to the verification region at 0 days. The targets do not correspond to the same meteorological system throughout the period.

A separate target becomes evident over eastern China (28°N, 120°E) at −5 days (labeled C). In a similar manner to A, this target can be traced downstream to −2 days, although with minimal baroclinic energy conversion and EKE in the location. Target C rapidly approaches target A at lead times shorter than −2 days to produce an elongated target region along the ridge in a broad area of EKE. Therefore, target C can arguably be traced backward from North America to eastern China over a 5-day period.[4] We also note that a fourth target area appears at −3 days over eastern China (30°N, 120°E), labeled D. However, it is questionable whether perturbations would propagate directly from this location to the verification region within 3 days. Instead, it is more likely that the

---

[4] We note that the targets were labeled subjectively, and may be ambiguous particularly when they are elongated, as the targeted observing time approaches the verification time.

broad error covariance structure in the ETKF has created spurious signal variance far downstream of China, which is in turn dynamically connected over 3 days with North America. Similar arguments may be applied to targets downstream of North America. It remains an open question whether perturbations in the North Atlantic storm track are associated with "upstream development" and act to modify forecasts over North America.

Returning to the targets labeled B, these targets would be the most perplexing to the forecaster or mission coordinator who is responsible for the decision on deployment. On examining -7 or -6-day lead times alone, one would suggest that B is a genuine target. And it may remain a genuine target at later times, although other target areas such as A and later C become a much higher priority. From the evidence presented here, one cannot conclude whether B is genuine or spurious, and the authors would recommend that A and C are safer options because of their connection to the verification region.

The longitudinal evolution of the targets as the lead time decreases from -7 to 0 days is summarized in a Hovmöller diagram (Fig. 3a). By finding the maximum value of the ETKF signal variance over a $20°$–$70°$ latitude band for each longitude, and repeating the process for each forecast lead time (in 12-h increments), a time–longitude plot is created. While there is subjectivity in how the Hovmöller diagram is prepared, the method employed here exhibits the highest consistency with the individual plots in Fig. 2. Figures 2 and 3a suggest that target A shifts eastward by $12°$ longitude day^{-1}, while C shifts eastward by $20°$ day^{-1}, prior to the -2-day lead time. Between -2 and 0 days, the targets shift much farther eastward, with C changing by approximately $40°$ longitude day^{-1}. We emphasize that the Hovmöller diagram illustrating the ETKF targets be interpreted with caution, as it does not strictly represent propagation of a physical quantity, but rather the continuity (or lack thereof) of targets at discrete times and whether they appear to have a dynamical connection with the verification region.

To examine whether the targets are associated with upper-tropospheric wave packets, a Hovmöller diagram of the 300-hPa meridional wind perturbation (about the monthly mean) at $45°$N is produced. Using the method of Zimin et al. (2003), the envelope for zonal wavenumber range of 4–11 is extracted, revealing a wave packet that propagates into the verification region with a group velocity of $\sim 20°$ day^{-1} (Figs. 3b,c). Target A follows the trailing edge of the wave packet from -6 to 0 days. At earlier times (-6 to -3 days), when it is coincident with an area of baroclinic energy conversion on the eastern side of the trough (Figs. 2b–d), target A follows the phase velocity of the wave crest in Fig. 3b. As the target time approaches

the verification time, A is more consistent with the group velocity (Fig. 3c). Target C shifts eastward with a speed similar to the group velocity, although it is in a location of minimum wave packet amplitude.

b. Case 13: Downstream baroclinic development

Case 15 did not exhibit a signature of downstream baroclinic development in the area where the targets shifted eastward. A more distinct association was evident in case 13. Between -3 and -2 days, the optimal target location shifts from an upstream location of high EKE (around $170°$W; Fig. 4c) to a downstream location (around $130°$W; Fig. 4a), which had been a secondary target. This may be interpreted as a jump of $40°$ longitude over 1 day, also shown in Fig. 4g. A significant value of ageostrophic geopotential flux convergence (diamonds in Fig. 4g, circles in Figs. 4d–f) is propagating downstream, amplifying the EKE in the downstream location that is also coincident with the ETKF target that dominates at -2 days and is connected to the verification region.

4. General properties of targets

a. Summary of 20 cases

The case examinations in section 3 revealed some preliminary conclusions. First, there existed optimal target areas in case 15 that could be traced as far back as Japan (target A, -7 days) and even eastern China (target C, -5 days). Second, the targets sometimes existed in locations of high baroclinic energy conversion upstream and eddy kinetic energy that amplified via downstream baroclinic development. Multiple target areas were present, increasing the ambiguity about where to deploy observations. This ambiguity may be partially resolved via the identification of spatiotemporal continuity of the targets in Hovmöller diagrams, which are presented for all 20 cases in Fig. 5 and summarized in Table 2. The individual ETKF guidance maps and EKE diagnostics at each observing time were also examined in detail but are omitted here for brevity.

A cursory examination of Fig. 5 indicates a range of "speeds" at which the target areas approach the verification region as the lead time decreases. First, targets can sometimes be found far upstream, near Japan ($140°$E) at 4–7-day lead with a subjective judgment of a connection to the verification region at later times (cases 4, 5, 6, 11, 12, 13, 14, 15, 16, 19, and 20). In half of these cases, a target is even discernible over eastern China ($120°$E). The targets near Japan are usually situated in a region of high baroclinic energy conversion. At later times (normally 1–3-day lead), the corresponding targets are sometimes associated with an area of significant ageostrophic geopotential flux convergence, in which the downstream
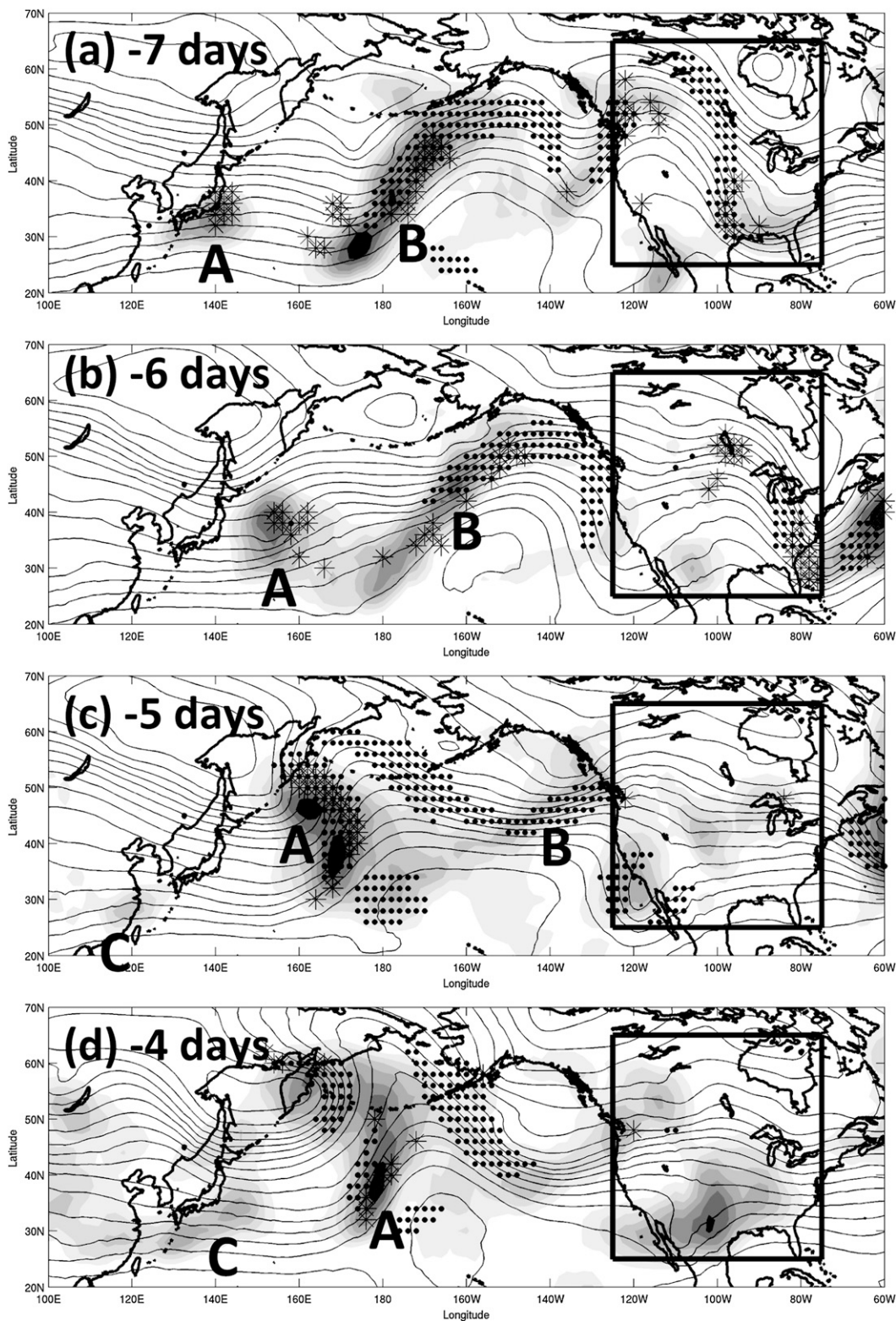
FIG. 2. Case 15. ETKF summary map of total energy signal variance for lead times from −7 days to 0 days, for a rectangular verification region over North America (shading). The ETKF values are normalized over the domain shown. Eddy kinetic energy (dots) and baroclinic energy conversion (asterisks). 500-hPa geopotential height (contours).
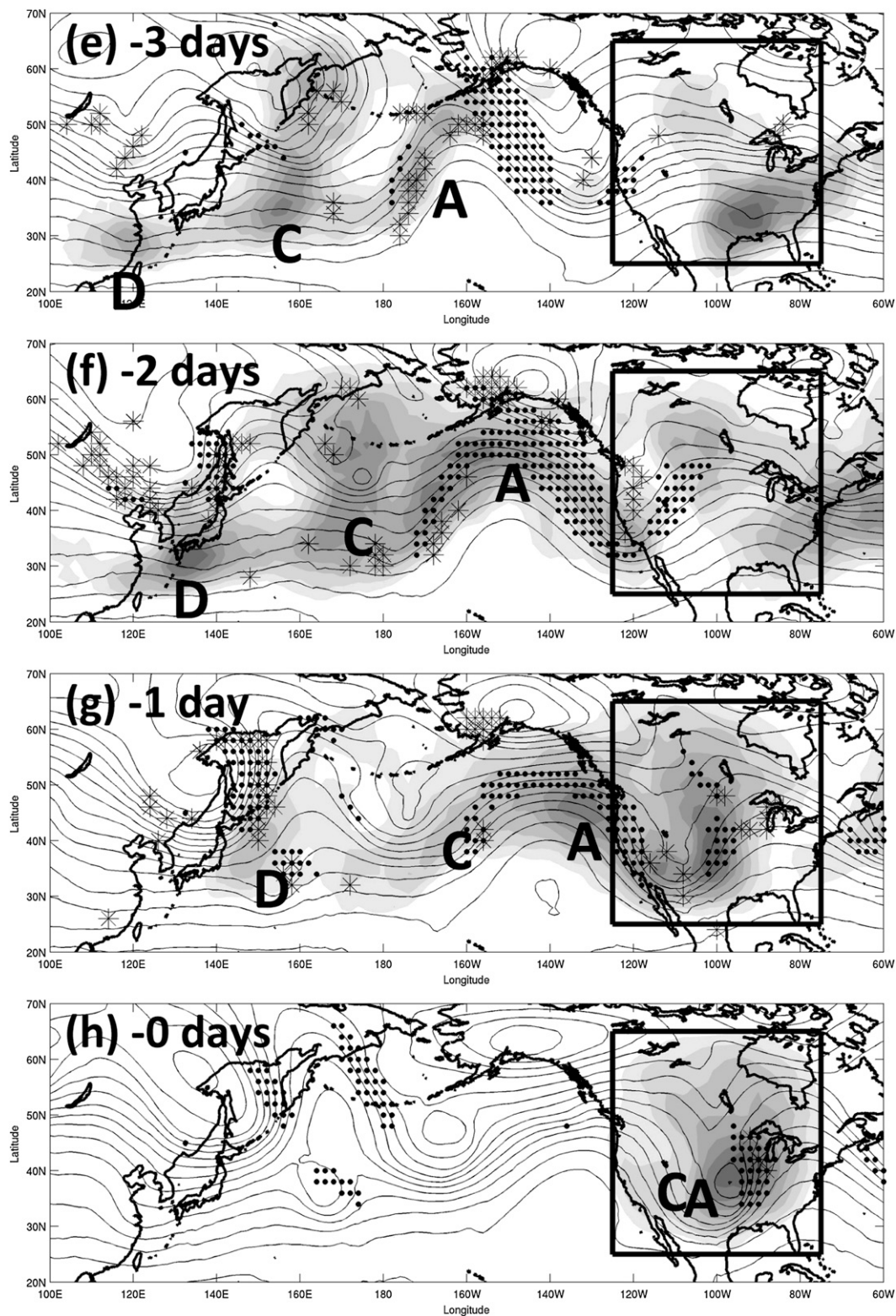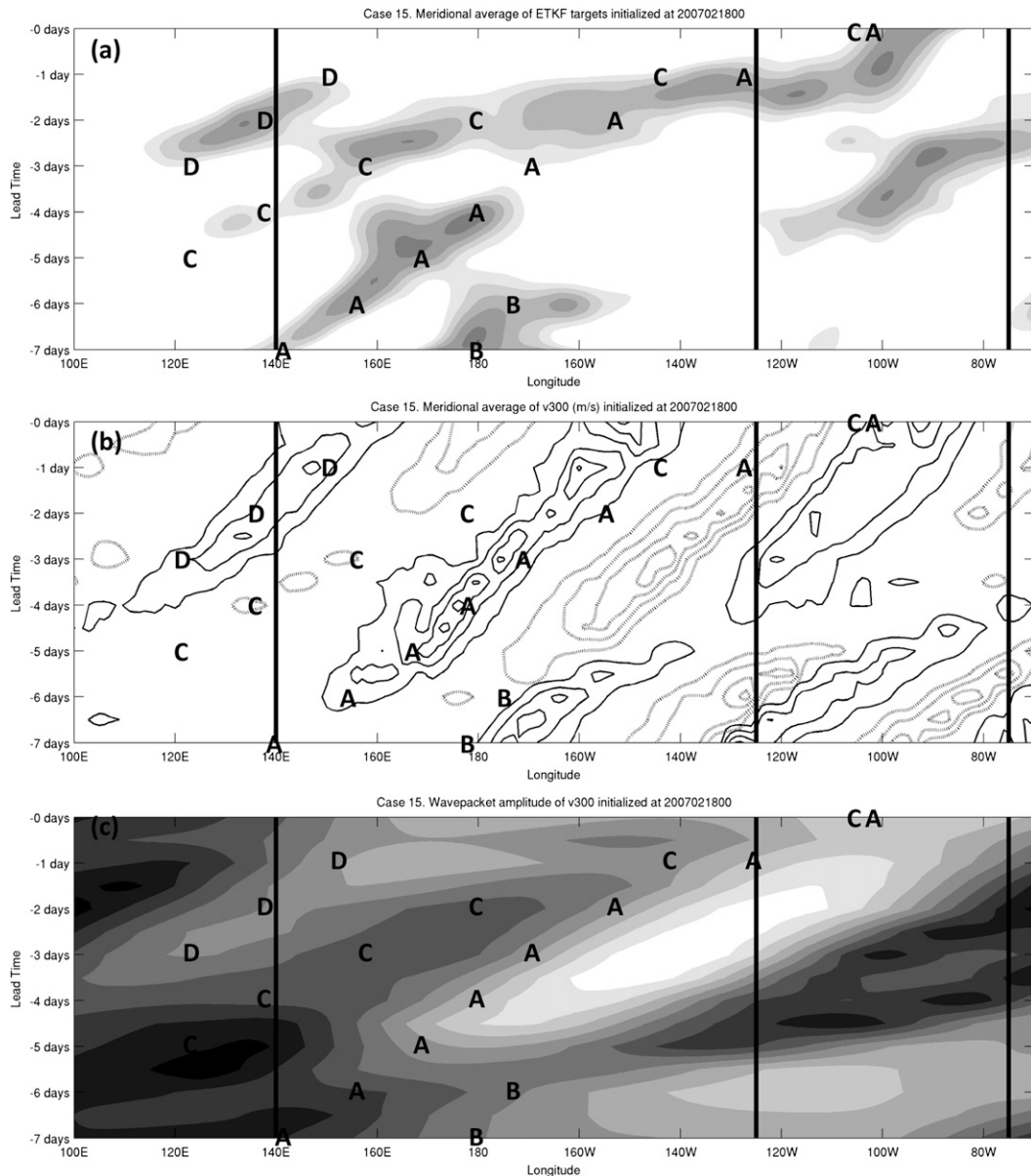
FIG. 2. (*Continued*)

FIG. 3. (a) Hovmöller diagram of ETKF targets for case 15, from 0- to −7-day lead time. (b) Hovmöller diagram for NCEP GFS forecast of 300-hPa meridional wind anomaly, initialized at 0000 UTC 18 Feb 2007. Positive anomaly (solid contours) and negative anomaly (dotted contours). (c) Amplitude of corresponding wave packet envelope.

system is receiving additional eddy energy (4, 6, 12, 13, and 14). In several cases, the targets downstream correspond to maxima of EKE, particularly at lead times of 2–4 days. Second, there exist slow-moving targets that linger over the central Pacific (date line) at lead times of 5–7 days (cases 3, 6, 7, 8, 9, 11, and 12). These cases are often associated with blocked flows over the Pacific Ocean. This quasi-stationarity is also consistent with studies of signal propagation produced by the NCEP GFS (Szunyogh et al. 2002; Hodyss and Majumdar 2007; Sellwood et al. 2008), in which parts of the signal were

found to propagate very slowly. Overall, for most cases, there is a discernible target that extends at least as far west as 160°E and can be traced toward the verification region.

While the targets often appear to evolve continuously toward the verification region as the lead time decreases, there are times in which a distinct shift in targets appears, as in Fig. 4. This shift represents an upstream target that was initially favorable for adaptive sampling at earlier times, before becoming relatively unimportant as a separate synoptic system downstream became the dominant
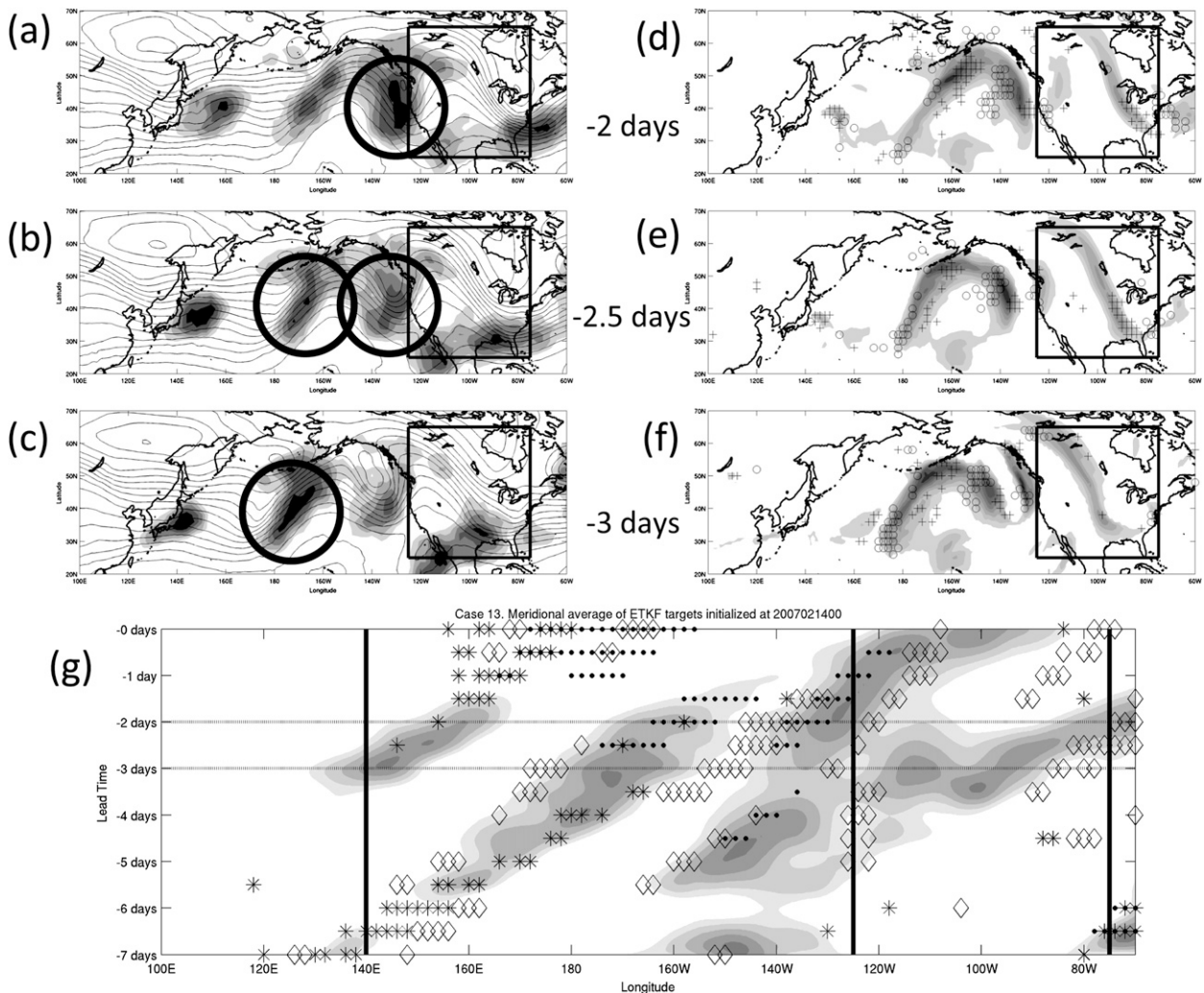
FIG. 4. (a)–(c) (shading) ETKF summary maps of signal variance for lead times of −2, −2.5, and −3 days, for case 13. (contours) NCEP GFS forecasts of 500-hPa geopotential height. (circles) Optimal target regions discussed in text. ETKF values are normalized with respect to the highest value on each map. (d)–(f) Corresponding NCEP GFS forecasts of eddy kinetic energy (shaded) and ageostrophic geopotential flux divergence [plus (+) = divergence, open circle (○) = convergence]. (g) Hovmöller diagram illustrating evolution of ETKF guidance (shading), and locations of significant baroclinic energy conversion (asterisks), eddy kinetic energy (dots), and ageostrophic geopotential flux convergence (diamonds). ETKF signal variance is normalized with respect to its maximum at each forecast time. The eddy diagnostics are normalized with respect to their maximum value on the Hovmöller diagram. The horizontal dotted lines indicate lead times of −2 and −3 days, which correspond to (a)–(f).

target. One may be tempted to link the jump in optimal target areas with downstream baroclinic development, in which centers of EKE similarly weaken upstream and strengthen in a separate system downstream. However, while this notion may sometimes hold (e.g., case 13), it is not always the case. We instead speculate that the shifting in targets is common for dispersive waves regardless of downstream baroclinic development. The targets may be modified because of the representation of the routine observational network, and they may sometimes be due to spurious long-distance correlations in the analysis. We will illustrate this in section 4b with a simple Rossby

wave example. It is also worth noting recent relevant work on propagation speeds in the midlatitudes. For example, the average group velocity of Rossby wave packets is approximately $30° \, \text{day}^{-1}$ (Szunyogh et al. 2002), and signal propagation may follow this speed (such as in Fig. 3b). Hakim (2003) had concluded from an observational analysis that the leading edge of a wave packet crosses North America at the tropopause around 3 days after the incipient disturbance was situated over the coast of eastern Asia. And in an investigation of propagating analysis and forecast errors over the northern Pacific Ocean, Hakim (2005) demonstrated that the group speed

of the errors is close to that of the mean zonal flow. Finally, parcel velocities in the midlatitude storm track may be as high as $60°$ day^{-1}, based on forward integrations of NOAA's Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Draxler and Rolph 2003). The shift in the targets with time could therefore be associated with the group velocity of Rossby waves, the mean zonal flow, or advective speeds, combined with the statistical correlation with the verification region at the verification time.

A coherent wave packet, determined using the 300-hPa meridional wind perturbation in the GFS forecast, is evident during two periods: one short period that spans cases 4 and 5 and a week-long period that includes cases 12, 13, 14, 15, and 16 (e.g., Fig. 3c). Interestingly, a target can be traced back to Japan in all of these cases in which a wave packet exists. However, these targets are usually not coincident with the location of highest wave packet amplitude, although there is high EKE. We also investigate the zonality, estimated as the difference between the 500-hPa zonal geostrophic wind in the NCEP GFS forecast and the monthly average zonal geostrophic wind at $45°$N (Horel 1985; Sellwood et al. 2008). For case 8, a blocking regime existed in the central and eastern Pacific, and little spatiotemporal continuity of the target areas was evident (Fig. 6a). In contrast, a strongly zonal regime existed for case 15 (Fig. 6b), which was examined in section 3. To provide quantitative estimates of the zonality associated with the target areas, spatial and temporal averages of the geostrophic wind anomaly over the western ($120°$–$160°$E, 4–7-day lead), central ($160°$E–$160°$W, 2–4-day lead), and eastern ($160°$–$120°$W, 0–2-day lead) Pacific were computed at $45°$N for each of the 20 cases. The cases corresponding to the six highest (most positive zonal anomaly) values and six lowest (most negative zonal anomaly) values are indicated by plus ($+$) and minus ($-$), respectively, in Table 2. In many of those cases in which the targets could be traced back to $140°$E, the flow was significantly zonal (5, 13, 14, 15, 16, and 19). And the cases in which a plausible target far upstream was not evident corresponded to regimes of highest negative zonal anomaly (1, 2, 7, 8, 10, and 11). It is also evident from Table 2 that the strongest associations between continuity of targets and zonality exist for the $160°$E–$160°$W region, which is a commonly used area in investigations of atmospheric blocking. Finally, it is worth noting that in nonzonal flows, the spatiotemporal continuity of the targets is not only less evident than in zonal flows as illustrated here, but according to Sellwood et al. (2008) the ETKF is also significantly less reliable in these flows.
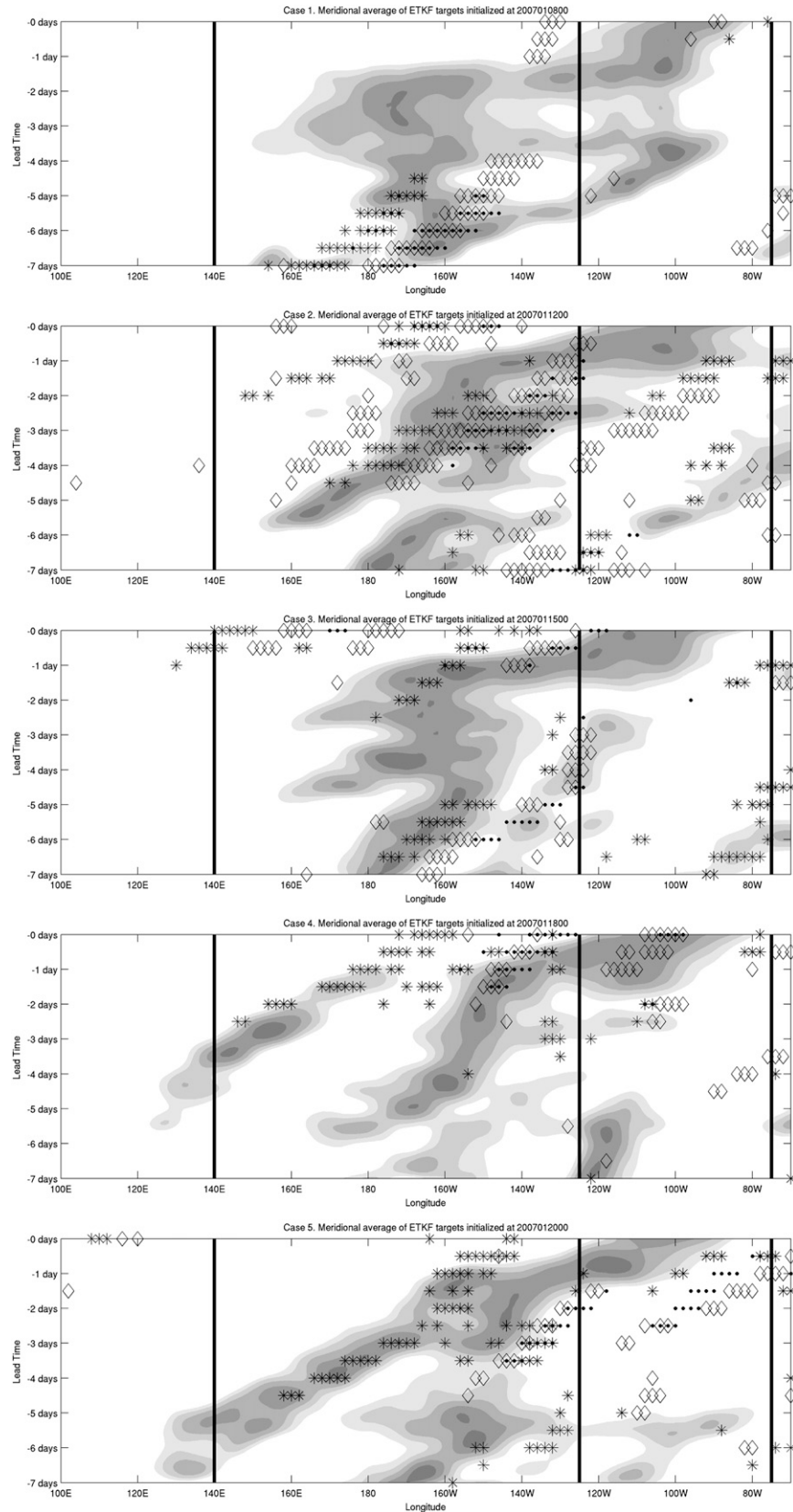
The primary target areas lie almost exclusively within the midlatitude storm track. Little sensitivity was found north of $55°$N in the northwestern Pacific. In one of the later cases (case 19, in mid-March), sensitivity was exhibited in the tropics, with a clear propagation of the target area toward the verification region. Given that the ETKF commonly selects areas of large ensemble spread in its targets, it may not emphasize the sensitivity captured by initially tiny convective perturbations that modify a forecast in the midlatitudes (Zhang et al. 2003; Hodyss and Majumdar 2007). The properties of the targets were also examined for different observing variables. The evolution of the targets was ambiguous in most cases. For example, although 850-hPa T is expected to be a useful variable for targeting far upstream, given the high baroclinicity near Japan, it is impossible to track sensitive areas across the Pacific because of low gradients in 850-hPa T. It was found that the most coherent targets were obtained by using a combination of u, v, and T at 200, 500, and 850 hPa, respectively, capturing the sensitivity in both low-level baroclinic zones and in jet regions aloft, and we retained this norm throughout the paper.

In summary, some common signatures were found in the evolving targets. In several cases (4–6 and 12–16), the targets could be traced back to Japan, a distinct wave packet existed, the flow was predominantly zonal, and there was high baroclinic energy conversion and EKE associated with the targets. In cases 12–16, downstream baroclinic development was judged to be occurring over the central Pacific. In contrast, there existed other cases (1, 7–10, and 17–18) where the targets were disjointed in space and time. The flow was often blocked in these cases, and/or the EKE diagnostics were less clear.

b. Discussion on continuity of targets

The targets generally shift steadily eastward ($10°$–$15°$ day^{-1}), until they reach the central Pacific. A shift on the order of $30°$–$50°$ day^{-1} in an area of rapid flow in the jet stream toward the verification region is then observed at

$\longrightarrow$

Fig. 5. Hovmöller diagrams summarizing 0–7-day evolution of ETKF targets (shaded), for all 20 cases. Also shown are baroclinic energy conversion (asterisks), eddy kinetic energy (dots), and ageostrophic geopotential flux convergence (diamonds). Values of ETKF signal variance are normalized with respect to its maximum at each forecast time. Values of the eddy diagnostics are normalized with respect to the maximum value on each Hovmöller diagram.
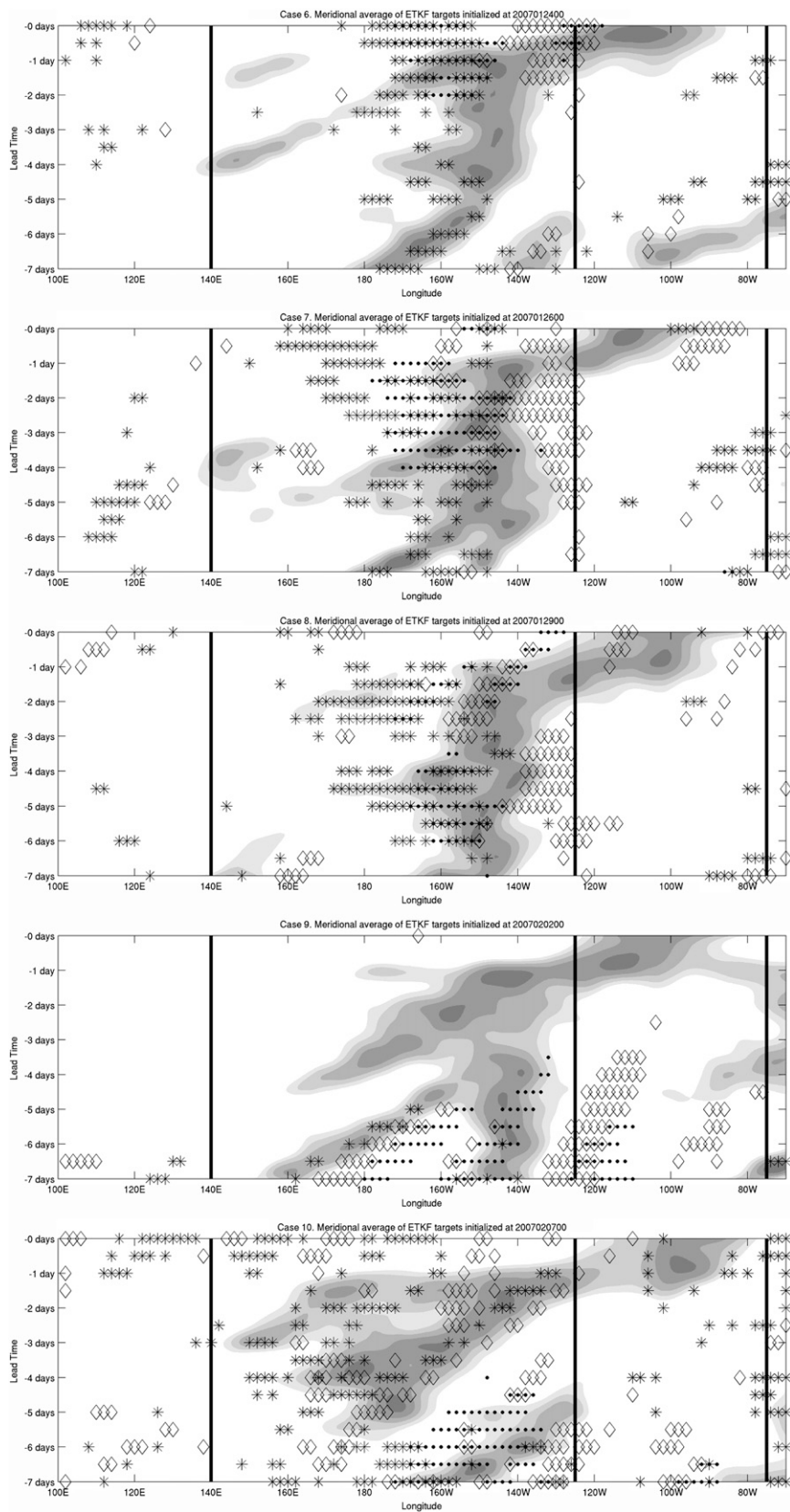
Case 1. Meridional average of ETKF targets initialized at 2007010800

Case 2. Meridional average of ETKF targets initialized at 2007011200

Case 3. Meridional average of ETKF targets initialized at 2007011500

Case 4. Meridional average of ETKF targets initialized at 2007011800

Case 5. Meridional average of ETKF targets initialized at 2007012000
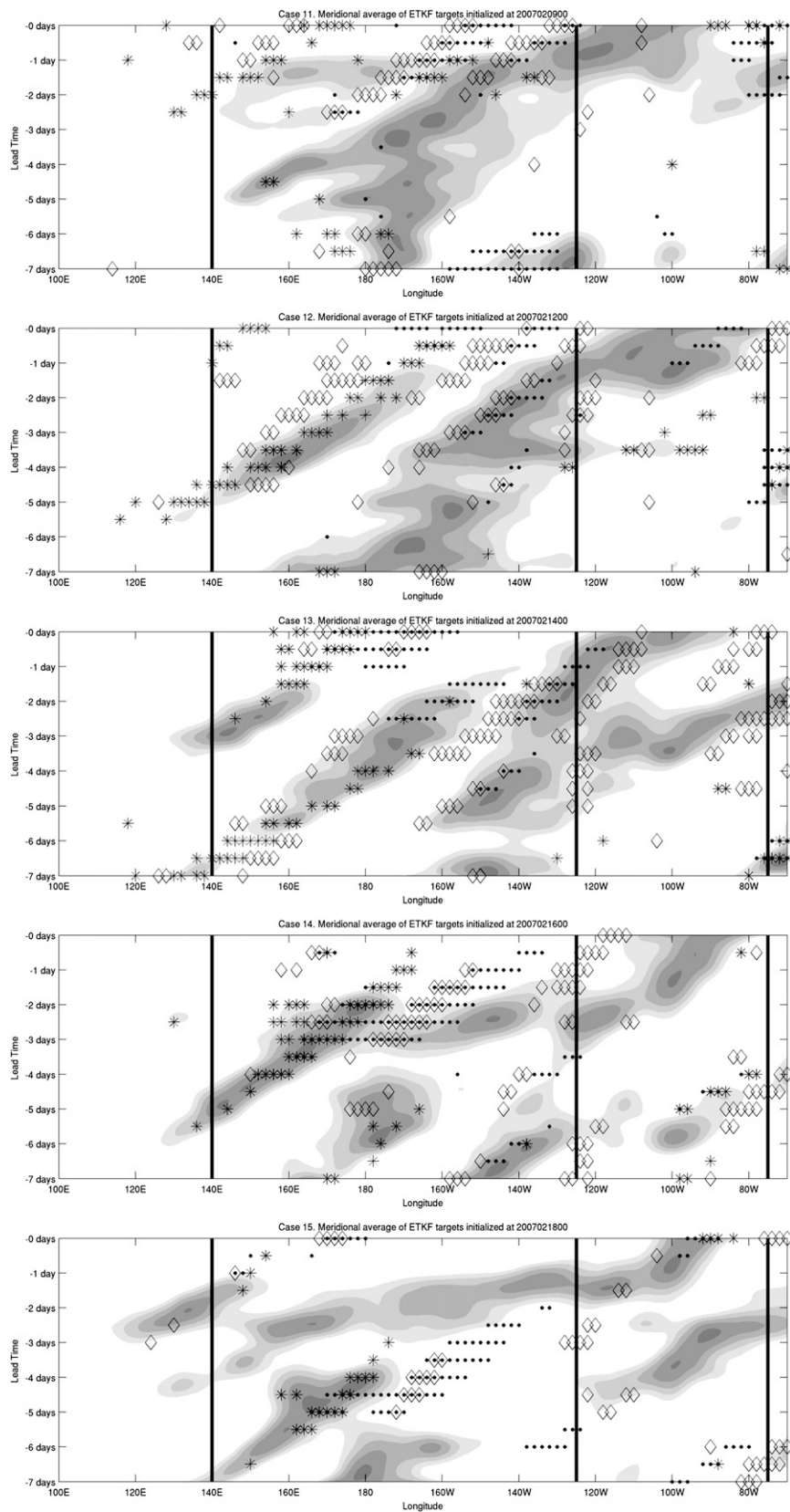
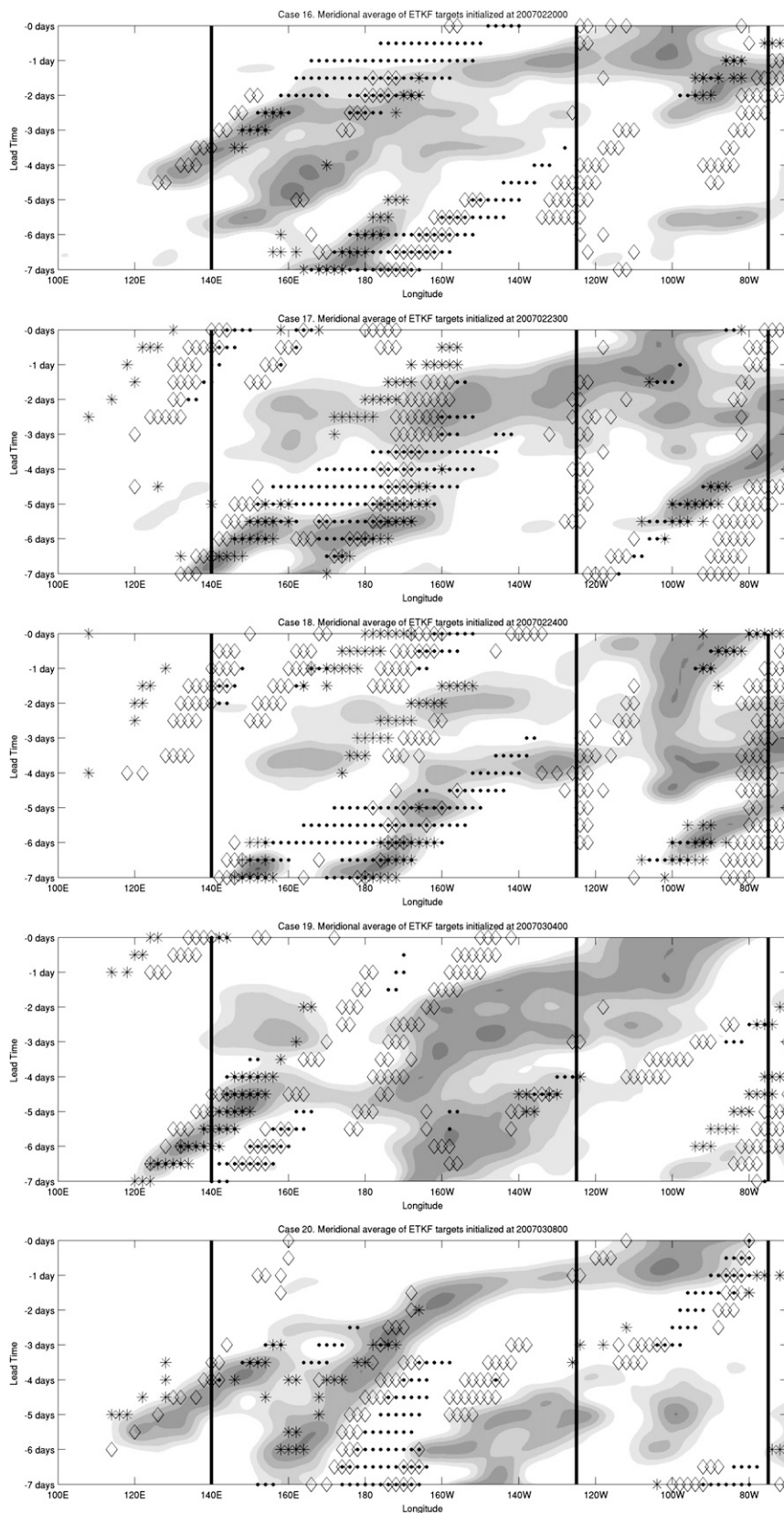FIG. 5. (*Continued*)

FIG. 5. (*Continued*)

FIG. 5. (*Continued*)

TABLE 2. Summary of ETKF results and diagnostics for all 20 cases. Cases in bold exhibit the most clearly discernible ETKF targets back to Japan. Cases in italics exhibit those with the weakest spatiotemporal coherence of targets. Integers in columns 2–4 correspond to lead time (in days) at which the targets lie in similar locations to maxima in eddy diagnostics [eddy kinetic energy (EKE), baroclinic energy conversion (Barcl), ageostrophic geopotential flux convergence (Ageo)]. Numbers in columns 5–7 give the forecast lead time (if any) at which the target is traced back to 140°E, 160°E and 180°. The plus (+) and minus (−) signs correspond to the 6 most zonal and the 6 least zonal flows of the 20 cases, for the western (4–7 day), central (2–4 day) and eastern (0–2 day) Pacific Ocean regions at the listed lead times. The existence of a discernible wave packet (WP) is also indicated (Yes/No).

	EKE	Barcl	Ageo	140°E	160°E	180°	W 4–7	C 2–4	E 0–2	WP
1		*5, 4*			*3*		−		−	*N*
2		5, 4, 3, 2	4, 3, 2, 1		6		−		−	N
3	1	7, 6, 5, 2, 1	1		5, 3	7			+	N
4	**1**	**3, 2, 1**	**1**	**4**	**6**				−	**Y**
5		**5, 4, 3, 2, 1**		**5**			+	+		**Y**
6	**2, 1**	**7, 6, 5, 3, 2**	**2, 1**	**4**		**7**	−			**N**
7	4, 3, 2	7, 6, 5, 4, 3	4, 3, 2		5	7		−	−	N
8	*7, 6, 5, 4, 2*	*6, 5, 4*	*3, 2, 1*			5		−	−	*N*
9	7, 6, 5	7, 6, 5	6		7, 5		+		−	N
10		5, 4, 3, 2, 1	5, 4, 3, 2, 1		6		−	−		N
11	2, 1		2, 1	5		7	−	−		N
12	**3, 2, 1**	**6, 5, 4, 3, 2**	**3, 2, 1**	**5**		**7**	−		+	**Y**
13	**3, 2, 1**	**7, 6, 5, 4, 3, 2**	**3, 2, 1**	**7**				+	+	**Y**
14	**3, 2**	**6, 5, 4, 3, 2**	**3, 2**	**5**	**7**		+	+		**Y**
15	**5, 4**	**7, 6, 5, 4, 3**	**2**	**4, 7**				+	+	**Y**
16	**3, 2**	**4, 3, 2**	**3, 2**	**4, 6**				+		**Y**
17	*6, 5, 4*	*7, 6, 5*	*3, 2*		*7*		+	−	+	*N*
18	*7, 6, 5*	*4, 3, 2*			*4*		+	−	+	*N*
19	7, 6, 5, 4	7, 6, 5, 4	5, 4, 3, 2	6			+	+		N
20	4, 3	6, 5, 4, 3, 2	5, 4, 3, 2, 1	5	7					N

lead times of 2 days or fewer for all cases except for case 5, and the target area is normally broad. During WSR, these short-range lead times are the norm, but with an ensemble that is (correctly) the most recently initialized at the time of computation. In contrast, the computations in this paper rely on an ensemble that is initialized 10 days prior to verification, based on the premise that a decision on continuous, medium-range targeting must be made at a long lead time. The drawback of this approach is that the ensemble forecast perturbations are 8–9 days old for the 1–2-day lead-time computations and therefore widely spread. The error covariance between remote locations in the storm track is thus likely to be high, yielding a broad spatial structure of an analysis increment based on an observation in the storm track, and an associated broad ETKF summary map structure. This increases the difficulty in interpreting the continuity of the target areas, and it is a challenge that the forecaster or field program planner will have to face in real time if the planning is necessary over a week prior to the weather event affecting society.

To provide clarification on the "shifts" in optimal targets as the lead time evolves, and the spurious targets that occur when the ensemble size is limited, ETKF guidance is computed for a simple ensemble of growing free barotropic Rossby wave perturbations (Holton 2004).

The ith perturbation $\Psi_i(x, y, t)$ comprises a linear combination of waves:

$$\Psi_i(x, y, t) = \sum a_j \psi'_j(x, y, t) \exp(\alpha t), \quad (8)$$

where $\psi'_j(x, y, t) = \text{Re}\{\exp[i(k_j x + ly - vt)]\}$ obeys the familiar dispersion relation $v = uk - \beta k/(k^2 + l^2)$ for a range of wavenumbers $k_j = [4, 11]$ and wavenumber $l = 1$, and mean zonal wind $u = 20$ m s^{-1} at 45°N. The wave growth rate is $\alpha = t/240$, for time t in hours. The variables in the state vector are arbitrary and are taken as dimensionless here without loss of generality (e.g., it could be a nondimensional streamfunction). The observation error covariance matrix is diagonal with variance values of 1. Point observations of the same dimensionless quantity are taken at every grid point in the two-dimensional space to provide the ETKF summary map guidance, with a verification norm of forecast error variance averaged within a fixed rectangular verification region.

Ensembles of two sizes are chosen: 400, to represent the full dimension of the physical space, and a reduced ensemble of 11 members. For the full ensemble, the ETKF targets are narrow and coherent, propagating toward the verification region as the lead time decreases to zero (Figs. 7a–f). More than one optimal target may exist (Figs. 7b,e). Furthermore, there is a distinct jump in the
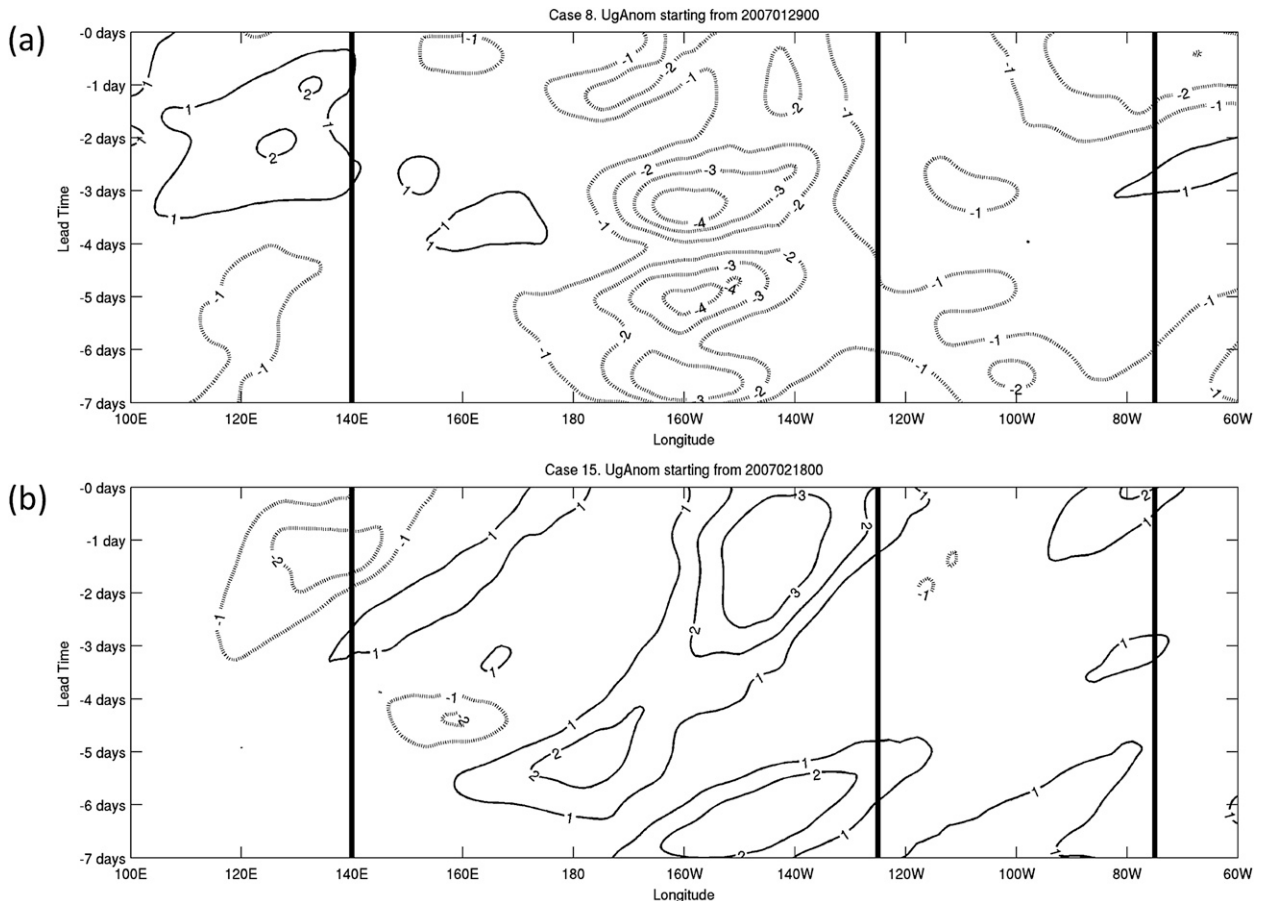
FIG. 6. (a) Hovmöller time series of geostrophic wind anomaly at 45°N, for case 8. (b) As in (a), but for case 15. High positive values indicate strongly zonal flow, and low negative values indicate weakly zonal or blocked flow.

optimal target similar to that of Figs. 4a–c, for example, between Figs. 7c and 7d. This jump is not due to spurious correlations given that the ensemble is of full rank, although nonlocal correlations do exist. For the reduced ensemble, a larger number of targets exists, and the main target in Fig. 7d is downstream of that for the full ensemble. Therefore, the spurious long-distance correlations have led to a different solution for the optimal target. The corresponding ensemble spread in Figs. 7g–l demonstrate that the ETKF guidance is not obviously associated with locations of large spread for the full ensemble, but a closer association with high spread exists in the reduced ensemble. Finally, the Hovmöller diagram in Fig. 7m demonstrates the discrete shifting targets, and their propagation with both the phase velocity ($\sim$7 m s^{-1}) and the group velocity ($\sim$30 m s^{-1}).

5. Conclusions

This paper documents a first attempt to investigate the nature of "target" locations identified by a modified version of the ETKF to sample tropospheric wind and temperature, to reduce medium-range forecast errors over North America. Twenty cases of potential high-impact weather during the winter of 2007 were selected. Using a 145-member ensemble comprising perturbations from NCEP, ECMWF, and CMC, ETKF summary maps within the verification region over North America were computed for lead times (verification time minus observing time) of 0–7 days. The following questions were examined:

1) Do the optimal target areas exhibit spatiotemporal continuity with lead time?
2) Is there a dynamical explanation for the properties of the target areas?

For the first question, the identification of targets that could be traced upstream over 12-h lead-time intervals was necessarily subjective. Nearly all cases exhibited at least one target that could be followed from the verification region at 0-day lead to the western Pacific Ocean (around 160°E) at 3–7-day lead. There were also multiple targets that did not exhibit such continuity. Perhaps the
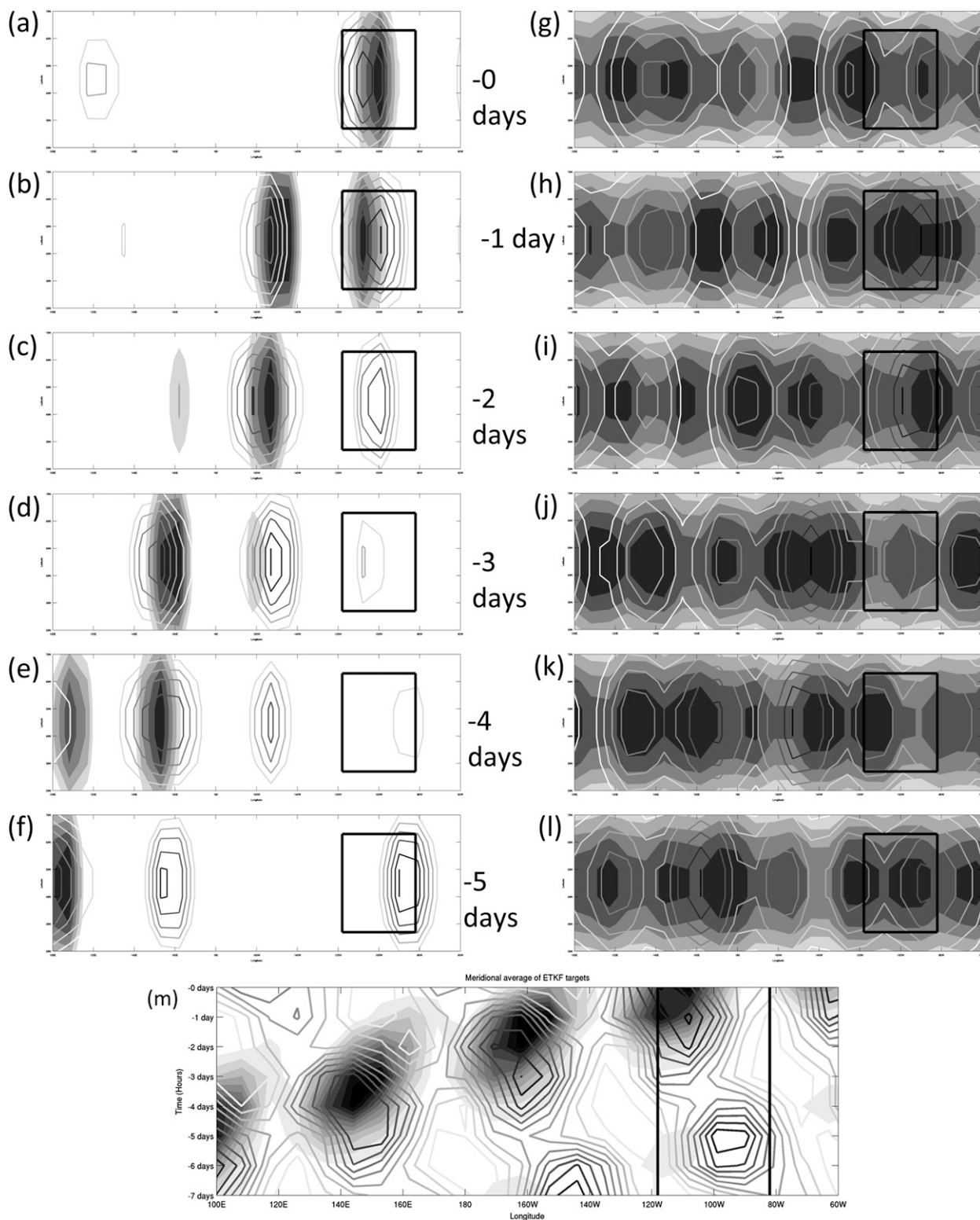
FIG. 7. (a)–(f) ETKF guidance as in Fig. 2, but for ensembles of free barotropic Rossby waves on a β plane. (g)–(l) Ensemble variance valid at the corresponding times. (m) Corresponding Hovmöller diagram of ETKF guidance. 11-member ensemble (contours). 400-member ensemble (shading). The verification region is represented by the black box.

most striking result was that in 11 of the 20 cases, targets existed as far upstream as 140°E, normally just offshore of Japan, 4–7 days prior to the verification time (although these were often not primary targets). In five of these cases, a target existed at 120°E, near eastern China, within a 7-day lead time. Given the dense observational network over eastern Asia, the targets were minimal for longer lead times. Significant targets were also evident in the central Pacific at 5–7-day lead times, shifting slowly eastward as the lead time was reduced. Given that the dispersive nature of the atmosphere is captured in the spreading of ensemble perturbations, the multiple target areas at longer lead times is not unexpected. It is, however, challenging to discriminate between targets that are genuine, and those that are spurious. We speculate that in situations where the optimal target region shifts from an upstream weather system to a separate system downstream over a day, the targets are genuine if errors associated with the adjacent systems are truly correlated. However, the limited ensemble size leads to the production of spurious long-distance correlations, and thereby false targets upstream. The concern is acute when the ensemble is over a week old and the correlations are broad in area and small in amplitude. At the corresponding short lead times, the target bears a broad structure over the eastern Pacific, although much of it is associated with rapid flow in the midlatitude jet. In general, in the absence of an error covariance localization that can be propagated forward in time, the ETKF will inevitably produce spurious targets, rendering the decision on deployment one of careful subjectivity. This is in contrast to the objective use of the ETKF during WSR programs, in which the target is normally unambiguous for the short-range forecast.

The investigation of the second question yielded mixed conclusions. In general, the majority of the targets resided in the midlatitude storm track. In the cases in which distinct targets existed offshore of Japan (4–6 and 12–16), the targets usually coincided with an area of baroclinic energy conversion that propagated downstream with a developing surface cyclone until the target time was 2–3 days prior to the verification time. In a few of these cases, in which spatiotemporal continuity existed from Japan eastward, the downstream targets were associated with maxima of eddy kinetic energy and ageostrophic geopotential flux convergence. While it is incorrect to conclude that the targets were always associated with downstream baroclinic development, cases 4–6 and 12–16 corresponded to the only two periods during which a coherent Rossby wave packet envelope was evident over the northern Pacific, even though the targets were not located within the envelope. In these cases, the flow was also mostly zonal, and the ETKF had previously been found to be reliable in predicting signal variance in such flows (Sellwood et al.

2008). The propagation speed of the targets is usually plausible, based on the findings of Szunyogh et al. (2002) and Hakim (2003, 2005) on the propagation of wave packets and analysis and forecast errors. In the remaining cases in our sample, a dynamical mechanism associated with the targets could not be established. This was particularly true for blocked flows, and it is worth noting that the ETKF is generally not reliable in such flow regimes (Sellwood et al. 2008).

In situations where the ETKF guidance is deemed to be reliable with appreciable continuity, we hypothesize that targeting over the western Pacific is important in zonal flows at lead times beyond 3 days. In particular, we propose that it is useful to sample just east of Japan, normally beyond 4 days in advance of the weather event, but usually not beyond 7 days. Our hypotheses on the suitability of targeting require systematic evaluation via data denial (e.g., Buizza et al. 2007; Kelly et al. 2007), using satellite data and in situ observations from field campaigns such as the recently completed winter phase of the THORPEX Pacific Asian Regional Field Campaign (T-PARC). The extent to which medium-range predictions of winter weather can be improved by targeted observations is likely to depend on the inherent predictability of the weather system; the locations, frequency, and accuracy of the observations; errors in the numerical model; and the data assimilation. To expand on the latter, if one of the flavors of ensemble Kalman filter used in midlatitude weather research (e.g., Szunyogh et al. 2008; Torn and Hakim 2008; Whitaker et al. 2008) were used to assimilate the data, the ETKF guidance would be expected to provide more accurate guidance than if an operational data assimilation scheme employing a quasi-isotropic error covariance were used. Concerns also remain about the validity of the assumption of linearity in all adaptive sampling strategies, although Reynolds and Rosmond (2003) have demonstrated that a linear technique is able to capture nonlinear forecast corrections on the synoptic scale, even at 3 days. Investigation is required on the limitations of the ETKF in nonlinear flows. Additionally, the structure of the ensemble perturbations at long lead times is likely not realistic, and therefore new techniques are required to rescale these perturbations to produce an error covariance structure that replicates that of a data assimilation scheme. And as is evident from our results, a 145-member ensemble still invokes spurious target areas. Hence, new methods need to be developed to mitigate the propagation of spurious error covariance information, even for large ensembles. Further improvements to the ETKF may be made via a superior representation of the routine observational network, higher vertical resolution, verification region selection that is based on the flow dynamics (Bishop et al. 2006), and an

improved estimate of error covariance via the THORPEX Interactive Grand Global Ensemble (TIGGE).

Although this investigation was performed in the context of continuous sampling via satellites, high-altitude unmanned aircraft, or driftsonde balloons over the northern Pacific Ocean, the assumption that those continuous observations would be assimilated prior to the next sampling time was not made here. Future studies using *serial adaptive sampling* (Majumdar et al. 2002) would sharpen the conclusions on where to sample at a given time, assuming that all routine *and targeted* observations had been assimilated up to and including that time. While we have focused on midlatitude target areas, the feasibility of sampling in the polar and tropical latitudes to improve forecasts of midlatitude weather requires investigation. Additionally, the potential and effectiveness of the method for other midlatitude, high-impact, warm-season weather systems such as strong mesoscale convective systems can be investigated in concert with an ensemble-based data assimilation scheme (Meng and Zhang 2008a,b). Such severe weather events are less baroclinic in nature, which leads to different behavior related to its predictability (Zhang et al. 2006). In general, extensive quantitative studies of the respective benefits of current and future observing systems are required, via Observing System Experiments (OSEs) and Observing System Simulation Experiments (OSSEs). These studies should be conducted in concert with investigations on predictability and dynamical processes associated with synoptic-scale, midlatitude weather systems, which has been identified as an important research challenge for the THORPEX community.

REFERENCES

Aberson, S. D., 2003: Targeted observations to improve operational tropical cyclone track forecast guidance. *Mon. Wea. Rev.,* **131,** 1613–1628.

Ancell, B., and G. J. Hakim, 2007: Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Mon. Wea. Rev.,* **135,** 4117–4134.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.,* **129,** 420–436.

——, ——, and ——, 2006: Verification region selection and data assimilation for adaptive sampling. *Quart. J. Roy. Meteor. Soc.,* **132,** 915–933.

Buizza, R., and T. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.,* **52,** 1434–1456.

——, and A. Montani, 1999: Targeted observations using singular vectors. *J. Atmos. Sci.,* **56,** 2965–2985.

——, D. S. Richardson, and T. N. Palmer, 2003: Benefits of increased resolution in the ECMWF ensemble system and comparison with poor-man's ensembles. *Quart. J. Roy. Meteor. Soc.,* **129,** 1269–1288.

——, C. Cardinali, G. Kelly, and J.-N. Thépaut, 2007: The value of targeted observations. Part II: The value of observations taken in singular vectors-based target areas. *Quart. J. Roy. Meteor. Soc.,* **133,** 1817–1832.

Chang, E. K. M., 1993: Downstream development of baroclinic waves as inferred from regression analysis. *J. Atmos. Sci.,* **50,** 2038–2053.

——, and I. Orlanski, 1993: On the dynamics of a storm track. *J. Atmos. Sci.,* **50,** 999–1015.

——, and ——, 1994: On energy flux and group velocity of waves in baroclinic flows. *J. Atmos. Sci.,* **51,** 3823–3828.

Danielson, R. E., J. R. Gyakum, and D. N. Straub, 2004: Downstream baroclinic development among forty-one cold-season eastern North Pacific cyclones. *Atmos.–Ocean,* **42,** 235–250.

Draxler, R. R., and G. D. Rolph, cited 2003: HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory) Model. NOAA/ARL, Silver Spring, MD. [Available online at http://www.arl.noaa.gov/ready/hysplit4.html.]

Gelaro, R., T. Rosmond, and R. Daley, 2002: Singular vector calculations with an analysis error variance metric. *Mon. Wea. Rev.,* **130,** 1166–1186.

Hakim, G. J., 2003: Developing wave packets in the North Pacific storm track. *Mon. Wea. Rev.,* **131,** 2824–2837.

——, 2005: Vertical structure of midlatitude analysis and forecast errors. *Mon. Wea. Rev.,* **133,** 567–578.

Hodyss, D., and S. J. Majumdar, 2007: The contamination of "Data Impact" in global models by rapidly growing mesoscale instabilities. *Quart. J. Roy. Meteor. Soc.,* **133,** 1865–1875.

Holton, J. R., 2004: *An Introduction to Dynamic Meteorology.* 4th ed. Elsevier Academic Press, 535 pp.

Horel, J. D., 1985: Persistence of wintertime 500 mb height anomalies over the central Pacific. *Mon. Wea. Rev.,* **113,** 2043–2048.

Hoskins, B. J., and P. J. Valdes, 1990: On the existence of storm-tracks. *J. Atmos. Sci.,* **47,** 1854–1864.

Kelly, G., J.-N. Thépaut, R. Buizza, and C. Cardinali, 2007: The value of targeted observations. Part I: Data denial experiments for the Atlantic and the Pacific. *Quart. J. Roy. Meteor. Soc.,* **133,** 1803–1815.

Langland, R. H., 2005: Issues in targeted observing. *Quart. J. Roy. Meteor. Soc.,* **131,** 3409–3425.

Leutbecher, M., 2003: A reduced rank estimate of forecast error variance changes due to intermittent modifications of the observing network. *J. Atmos. Sci.,* **60,** 729–742.

Lorenz, E. N., and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.,* **55,** 399–414.

MacDonald, A. E., 2005: A global profiling system for improved weather and climate prediction. *Bull. Amer. Meteor. Soc.,* **86,** 1747–1764.

Magnusson, L., M. Leutbecher, and E. Källén, 2008: Comparison between singular vectors and breeding vectors as initial perturbations for the ECMWF ensemble prediction system. *Mon. Wea. Rev.,* **136,** 4092–4104.

Majumdar, S. J., C. H. Bishop, B. J. Etherton, I. Szunyogh, and Z. Toth, 2001: Can an ensemble transform Kalman filter predict the reduction in forecast error variance produced by targeted observations? *Quart. J. Roy. Meteor. Soc.,* **127,** 2803–2820.

——, ——, ——, and Z. Toth, 2002: Adaptive sampling with the ensemble transform Kalman filter. Part II: Field program implementation. *Mon. Wea. Rev.,* **130,** 1356–1369.

——, S. D. Aberson, C. H. Bishop, R. Buizza, M. S. Peng, and C. A. Reynolds, 2006: A comparison of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Wea. Rev.,* **134,** 2354–2372.

McLay, J. G., C. H. Bishop, and C. A. Reynolds, 2008: Evaluation of the ensemble transform analysis perturbation scheme at NRL. *Mon. Wea. Rev.,* **136,** 1093–1108.

Meng, Z., and F. Zhang, 2008a: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study. *Mon. Wea. Rev.,* **136,** 522–540.

——, and ——, 2008b: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part IV: Comparison with 3DVAR in a month-long experiment. *Mon. Wea. Rev.,* **136,** 3671–3682.

Namias, J., and P. F. Clapp, 1944: Studies of the motion and development of long waves in the westerlies. *J. Meteor.,* **1,** 57–77.

Orlanski, I., and J. Katzfey, 1991: The life cycle of a cyclone wave in the Southern Hemisphere. Part I: Eddy energy budget. *J. Atmos. Sci.,* **48,** 1972–1998.

——, and E. K. M. Chang, 1993: Ageostrophic geopotential fluxes in downstream and upstream development of baroclinic waves. *J. Atmos. Sci.,* **50,** 212–225.

——, and J. P. Sheldon, 1993: A case of downstream baroclinic development over western North America. *Mon. Wea. Rev.,* **121,** 2929–2950.

——, and ——, 1995: Stages in the energetics of baroclinic systems. *Tellus,* **47A,** 605–628.

Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.,* **55,** 633–653.

Pellerin, G., L. LeFaivre, P. L. Houtekamer, and C. Girard, 2003: Increasing the horizontal resolution of ensemble forecasts at CMC. *Nonlinear Processes Geophys.,* **10,** 463–468.

Petersen, G. N., S. J. Majumdar, and A. J. Thorpe, 2007: The properties of sensitive area predictions based on the ensemble

transform Kalman filter (ETKF). *Quart. J. Roy. Meteor. Soc.,* **133,** 697–710.

Rabier, F., and Coauthors, 2008: An update on THORPEX-related research in data assimilation and observing strategies. *Nonlinear Processes Geophys.,* **15,** 81–94.

Reynolds, C. A., and T. E. Rosmond, 2003: Nonlinear growth of singular-vector-based perturbations. *Quart. J. Roy. Meteor. Soc.,* **129,** 3059–3078.

——, M. S. Peng, S. J. Majumdar, S. D. Aberson, C. H. Bishop, and R. Buizza, 2007: Interpretation of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Wea. Rev.,* **135,** 4006–4029.

Sellwood, K. J., S. J. Majumdar, B. E. Mapes, and I. Szunyogh, 2008: Predicting the influence of observations on medium-range forecasts of atmospheric flow. *Quart. J. Roy. Meteor. Soc.,* **134,** 2011–2027.

Simmons, A. J., and B. J. Hoskins, 1979: The downstream and upstream development of unstable baroclinic waves. *J. Atmos. Sci.,* **36,** 1239–1254.

Szunyogh, I., Z. Toth, A. V. Zimin, S. J. Majumdar, and A. Persson, 2002: On the propagation of the effect of targeted observations: The 2000 Winter Storm Reconnaissance Program. *Mon. Wea. Rev.,* **130,** 2803–2820.

——, E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, E. Satterfield, and J. A. Yorke, 2008: A local ensemble transform Kalman filter data assimilation system for the NCEP global model. *Tellus,* **60,** 113–130.

Torn, R. D., and G. J. Hakim, 2008: Performance characteristics of a pseudo-operational ensemble Kalman filter. *Mon. Wea. Rev.,* **136,** 3947–3963.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.,* **125,** 3297–3319.

Velden, C., and Coauthors, 2005: Recent innovations in deriving tropospheric winds from meteorological satellites. *Bull. Amer. Meteor. Soc.,* **86,** 205–223.

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operation forecast system. *Tellus,* **60A,** 62–79.

Weissmann, M., R. Busen, A. Dörnbrack, S. Rahm, and O. Reitebuch, 2005: Targeted observations with an airborne wind lidar. *J. Atmos. Oceanic Technol.,* **22,** 1706–1719.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.,* **136,** 463–482.

Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.,* **60,** 1173–1185.

——, A. M. Odins, and J. W. Nielsen-Gammon, 2006: Mesoscale predictability of an extreme warm-season precipitation event. *Wea. Forecasting,* **21,** 149–166.

Zimin, A. V., I. Szunyogh, D. J. Patil, B. R. Hunt, and E. Ott, 2003: Extracting envelopes of Rossby wave packets. *Mon. Wea. Rev.,* **131,** 1011–1017.